DOCUMENT RESUME

ED 145 367 CS 003 677

AUTHOR Henry, G.; Grisay, A.
TITLE Methodological Aspects of Evaluation in Primary Reading.
PUB DATE [70]
NOTE 28p.; Study prepared at University of Liege, Belgium; Not available in hard copy due to marginal legibility of original document

EDRS PRICE MF-$0.83 Plus Postage. HC Not Available from EDRS.
DESCRIPTORS Beginning Reading; *Decoding (Reading); Models; Primary Education; *Reading Comprehension; *Reading Research; *Reading Skills; *Test Construction; *Test Items

ABSTRACT

This paper develops a model for generating sets of replicable items for testing a range of reading skills in the primary grades. The procedure is particularly concerned with tests to identify a child's profile in reading achievement and to inform a teacher, principal, or district of the actual level of achievement in reading. Although the model is more general, this application focuses on only two aspects of reading: coding and decoding skills and reading comprehension. (AA)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

METHODOLOGICAL ASPECTS OF EVALUATION IN PRIMARY READING

"PERMISSION TO REPRODUCE THIS MATERIAL BY MICROFICHE ONLY HAS BEEN GRANTED BY

G. Henry
A. Grisay

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND THE ERIC SYSTEM CONTRACTORS."

by G. HENRY and A. GRISAY
University of Liège (Belgium)

ED145367

For many decades, a lot of hard work has been devoted to research in primary reading. The study of methods, basic programs, and reading failures..., has been highly emphasized while more fundamental problems have been neglected.

In this context, the evaluation of primary reading achievement has been rather poorly approached.

J. Chall notes (1) that the comparison of methods (look-say, systematic phonics or intrinsic phonics) is very difficult and biased by the large diversity of the means used by the investigators when evaluating reading achievement. She compares, therefore, any available studies according to eight measures of reading ability.

1. Word pronunciation
2. Connected oral reading
3. Phonics
4. Spelling
5. Vocabulary
6. Silent reading comprehension
7. Rate of reading
8. Interest, fluency, expression.

---

(1) J. CHALL, 1967.

CS003677

BEST AVAILABLE COPY

However, it must be emphasized that labels such as "silent reading comprehension" are very imprecise. J. Chall again defines "silent reading comprehension" as "the ability to understand material read. Sometimes called paragraph reading, this is usually part of a standardized silent reading test. The child reads selections and answers questions to show that he understands them. Like the vocabulary subtest, the comprehension subtest measures an indeterminate combination of skills and is affected by rate of reading". It would be impossible to confess more explicitly the inability of the reading researcher to define clearly what he plans to measure.

We are convinced that no substantial progress can be expected in the domain of reading evaluation without a large amount of research devoted to theoretical problems.

Before proposing a tentative way for the evaluation in primary reading (1st and 2nd grades), let us list what seems to us the most important handicaps to any progress in the knowledge of primary reading:

a) In a selective system, the explicit (or implicit) objective of reading tests is to ascertain "which children we have to retain in the low grades". In a more democratic system, the reading tests help answering the question: "Which child needs remedial instruction in primary reading ?". Though the tests yeald accurate and reliable composite scores, they do not provide, excepted in a few cases, any further data about specific aptitudes or deficits of the child.

b) People interested in reading instruction are mostly teachers, principals, administrators, basic readers publishers, test publishers,... Few of them are trained in linguistics or are highly interested in theoretical problems related to reading evaluation.

c) No test explores the whole range of reading behaviors.

d) No criterion-referenced procedure is used for selecting or writing items; as a consequence, a too large amount of freedom is left to the test-writer.

In this paper, we suggest a new way for generating sets of replicable items for testing a wide range of reading behaviors. This procedure could be of considerable help for the researcher with following objectives:

a) Identifying a child's profile in reading achievement.

b) Informing a teacher, a principal, a district or board administrator of the actual level of achievement in reading instruction.

To begin with, we have focused our attention on two aspects of reading evaluation only: coding and decoding skills, and reading comprehension. Let us first consider the problems concerning the code; we shall next describe a model convenient for reading comprehension evaluation.

A. CODING AND DECODING SKILLS.

The model designed for evaluating the code aspects will be useful if most items (ideally all of the ...sed in standardized tests can be generated through a nearly automatized procedure and classified according to the dimensions of the model.

The first dimension refers to the media used by the test administrator for asking a question and by the student for answering it. Three categories (written question and written response, written question and oral response, oral question and written response) are included in the model.

The second dimension refers to the kind of operation performed by a student for completing the item. Two levels, recognition and production, may be required. In the case of "recognition" the subject has to select one or more correct responses among n decoys for matching a stimulus. In the case of "production" the subject has to construct one or more correct responses matching a stimulus.

The third dimension refers to the linguistic material used in any specific item. Five categories are related to this dimension: letters and sounds, syllables, words, sentences, paragraphs.

The whole model comprises 3 x 2 x 5 = 30 cells.

			Letters and sounds	Syllables	Words	Sentences	Paragraphs
Precoding	W Q — W R	Recognition	1	2	3	4	5
		Production	6	7	8	9	10
coding	W Q — O R	Recognition	11	12	13	14	15
		Production	16	17	18	19	20
	O Q — W R	Recognition	21	22	23	24	25
		Production	26	27	28	29	30

From this model, it proves very easy to derive the items connected to any cell. We must, however, bear in mind that such a model is useful if the items may be derived through a standardized criterion procedure. It must, moreover, be emphasized that the derivation procedures may vary according to the goals pursued by the test-writer. A review of the goals attainable through the model and of the means which could be used for that purpose will be found at the end of this paper.

Let us now consider which items correspond to each cell of the table and which construction problems we encounter.

1. A possible item derived from the cell number 1 would be:
"Choose from among four letters the same as the letter in front"

b a b c d

Two remarks about this item:

a) No genuine reading takes place. Without any exposure to reading teaching, a child could complete this item which could, however, be found in reading readiness tests.

b) A criterion procedure can be used for deriving these items by selecting a random sample of letters as stimuli. The decoys are also randomly selected from the 26 letters of our alphabet or from any sub-population of letters which offer some graphic characteristics. Visual discrimination studies could possibly indicate the best sub-population of letters usable as decoys.

2. Items corresponding to cells 2-3-4-5 can be derived the same way. The population of letters, sounds and syllables is however finite. The population of words, sentences or paragraphs is infinite, but it is quite possible to define finite sub-populations (Examples: words belonging to a basic vocabulary list, meaningless words created by adding given syllables or sounds, sentences drawn from a basic-reader,...).

As previously stated, such sub-populations have to be defined according to the objectives of the test-writer.

3. Items derived from cells number 6-10 aim at testing the ability to copy different materials.

4. In a more general way, items derived from cells number 1-5 and 6-10 do not call for genuine reading coding or decoding skills. No translation is made indeed from a written material to an oral one or conversely. Items derived from cells 1-5, moreover, do not require any learning of reading. Items derived from cells number 6-7 and in some cases from cell number 8 require only visual discrimination skills plus psychomotor skills, but no systematic learning of reading. We have included these items in the model as pre-coding items for at least two reasons:

   a) Though no genuine reading takes place, these items are, however, prerequisites for reading learning. The ability to cope with such an abstract material could possibly be a good predictor of reading achievement in the first grade.

   b) In fact, we call "reading skills" all the skills related in any way to printed linguistic material. Thus, we have not included in the model items combining oral questions and oral responses (Example: identify a spoken word among an oral choice). We are, however, aware that such items are very important as prerequisites for reading learning. They will be discussed hereafter, when we will consider more thoroughly items derived from cells 1-6-11-16-21-26.

5. Items corresponding to cells 11-15 may be derived by a similar procedure. A letter "p", for instance, is given as a written stimulus. The test administrator or a tape recorder utters sounds such as /p/, /a/, /b/, /u/. The student has to identify the correct sound.

6. Items corresponding to cells 16-20 measure the ability to read orally letters, syllables, paragraphs,...

7. Items derived from cells 21-25 are constructed according to the same procedure. These items are read by the test administrator. The student has to select the correct answer among written decoys.

8. Items derived from cells 26-30 measure the ability to transcribe letters, syllables,..., paragraphs spoken by the test administrator (dictation).

9. Cells 5-14-15-20 have been included in the model for comprehensiveness; we are aware that items derived from these cells are not very important and are even unrealistic.

So far, our approach to item derivation has been rather theoretical. Now, let us look at the problem of item construction and sampling from a more practical point of view. For instance, what are the problems to be solved for designing items corresponding to cells 1-6-11-16-21-26, i.e. all items related to letters and sounds (1).

---

(1) This part of the study has been conceived for the French Language. Discrepancies exist in this respect between French and English.

a) Items derived from cells 1 and 6 require mainly visual recognition of a graphic feature according to several clues (1).

- left or right-oriented ex. /
- single or double ex. /
- up or down oriented ex. /
- curled or not ex. /
- ...

b) Any stimulus or decoy relative to an item from cells 11-16-21-26 is characterized by phonological and graphic features. Phonological features of sounds have not been described heretofore for the model does not include items made up of oral questions and oral responses.

Before examining items derived from the cells 11-16-21-26, we have to deal with phonological features of sounds such as:

- voiced - non voiced ex. /t/ vs /d/
- stop - non stop ex. /p/ vs /f/
- bilabial - dental ex. /p/ vs /t/
- oral - nasal ex. /a/ vs /ã/
- ...

If we simultaneously take into account graphic properties of letters and phonological features of sounds, we understand easily that stimuli and decoys may be derived by a criterion procedure.

---

(1) The relative importance of these clues might be assessed by means of experimental studies. See, for instance, E.B. COLEMAN, 1970, and H.M. POPP, 1964.

For instance, sub-samples of items may randomly be derived from following sub-populations of letters and sounds and are representative samples.

1. Univocal correspondence

1.1. One sound corresponds to one letter. The correspondence is biunivocal.

Ex.: r, p, v, d

From these sounds, we may select well-identified stimuli and decoys.

1.1.1. The stimulus and the decoy(s) differ by one phonological clue while their graphic features differ by two clues at least.

Ex.:

Phonological clues	+ consonant	+ consonant
	+ stop	+ stop
	+ labial	+ labial
	non voiced	+ voiced
Graphic features	down	up
	non curled	curled
	right-oriented	right-oriented
	single	single

1.1.2. The stimulus and the decoy(s) differ by one graphic clue while their phonological features differ by two clues at least.

Ex.:

Phonological clues	voiced	non voiced
	non stop	stop
	consonant	consonant
	lateral	dental
Graphic features	up	up
	right-oriented	right-oriented
	single	single
	curled	non curled

1.1.3. The stimulus and the decoy(s) differ by one graphic and one phonological clue.

Ex.:

Phonological clues	consonant	consonant
	nasal	nasal
	voiced	voiced
	dental	bilabial
Graphic features	up	up
	single	double

1.2. A sound is written with a set of letters.

The correspondance can be:

- Biunivocal: in French the sound /u/ is always written ou, and ou is always spoken /u/
- Univocal (graphic): the graphic set of letters au is always spoken /o/ (but /o/ can be written as o, au or eau).
- Univocal (phonemic): the sound /ʃ/ is always written ch (but ch can be spoken /ʃ/ or /k/).

2. Multiple correspondance

All linguistic data to be tested in this domain may be described at best according to two dimensions:

- The direction of the decoding scheme
  a) W to O: a given letter or set of letters can be orally transcoded into different sounds.
  b) O to W: a given phoneme can be transcribed with different letters or sets of letters.
- The distribution of the allomorphs is either governed by contextual constraints or is conventional. In the latter case no rule does exist: the graphic or the phonological features of words have to be memorized.

It is, therefore, quite impossible to test multiple correspondences from sounds isolated from words. Multiple correspondences have to be tested within words. Items will thus aim at testing:

a) The ability of the student to identify the different decoding schemes.

b) The knowledge of the contextual constraints allowing to select one decoding scheme from among other possible schemes.

c) The memorizing of conventional graphic of phonological features.

Let us have a look to some possible linguistic data to be tested:

		Allomorphs are distributed according contextual constraints	Allomorphs are distributed conventionally
W→O	g	a, o, u → /g/ i, e → /ʒ/	ch → / / ex. (chat) → /k/ ex. (chrétien)
	s	between vowels → /z/ before consonants → /s/	x → /ks/ (axe) → /z/ (dixième) → /s/ (dix)
	c	a, o, u → /k/ e, i → /s/	
O→W	/ /	before bilabials /p/ or /b/ → m	/ã/ → en → an ex. pendant

B. READING COMPREHENSION

Let us refer again to the J. Chall's definition of "silent reading comprehension" quoted above (p. 2). It is confessed that "silent reading tests measure an indeterminate combination of skills". The entire responsibility for the situation cannot rely on testwriters only since no exhaustive theory of reading comprehension has been developed so far. When such a theory is available, the construct validity of reading tests will increase accordingly. In the meantime reading test-constructors can but depend upon existing theories.

If, for instance, we want to test the ability of a student to understand the content of a first or second grade passage, we construct four sub-samples of items.

1. Items measuring knowledge of the vocabulary used in the passage.
2. Items measuring mastery of the syntactical structures.
3. Items measuring ability to rely upon anaphoric relations.
4. Items measuring ability to use compensatory devices which supply the deficiencies of informations.

1. Vocabulary

First of all, we list the words which are absentees from a basic vocabulary list. For the French language, we use the CREDIF list (1). According to our purpose, we take into account, either the whole population of these words, or a randomized sub-sample of a given size.

---

(1) GOUGENHEIM, 1964.

In a second step, we read in the dictionary (1) the definition of every word selected (according to the specific use of the word in the passage).

Ex. Cime: "Highest part of a tree or a mountain".

From this definition, we drew the semic scheme of the word.

Ex. Cime: + extremity', + 'up', + 'tree'
'mountain'

In a third step, we look in a dictionary (2) providing synonyms and antonyms for words that complete the semantic space of the word under consideration.

---

(1) J. DUBOIS, 1967.

(2) R. BAILLY, 1971.

The following semic schema is then built.

	Semic kernel		Contextual semes			
	A (Extremity)	B (Up)	C (Tree)	D (Mountain)	E (Building)	F (Man)
CIME	+ (1)	+	+	+	o	o
SOMMET	+	+	+	+	+	o
FAITE (toit 2)	+	+	o	o	+	o
CRETE	+	+	o	+	o	o
TETE (2)	+	+	o	o	o	+
LE HAUT (2)	+	+	+	+	+	o
POINTE (2)	+	o/+	+	o	o/+	o
BOUT (2)	+	o/+	o	o	o	o
BORD (2)	+	o	o	o	o	o
LIMITE	+	o	o	o	o	o
FOND (2)	+	-	o	o	o	o
PIED (2)	+	-	+	+	+	+
BASE	+	-	+	+	+	o
LE BAS (2)	+	-	+	+	+	o

(1) + : the word is positively marked for a given seme ex. cime: + up.
- : the word is negatively marked for a given seme ex. fond: - up.
o : the word is neither positively neither negatively marked for a given seme, i.e., the seme does not belong to the semic scheme of the word.
o/+ : neutralized semes. The seme is unspecified for the neutral meaning of the word, but for the marked meanings it can be specified as +.

(2) Words belonging to the basic vocabulary.

From this table, one can easily derive items testing:

a) The semic kernel (seme B)

by selecting as decoys:

- all the words having a + for the seme A :

la cime | la tête ?
| le haut ?
| la pointe ?
| le bout ?
| le bord ?
| le fond ?
| le pied ?
| le bas ?

- or a sub-sample:

la cime | le haut ?
| le bord ?
| le bas ?

b) The contextual seme "tree" (seme C)

by selecting as decoys all the words having a + for semes A and B:

la cime = le toit, la tête, le haut, la pointe, le bout — d'un arbre

c) The contextual semes D-E-F

by simply manipulating the semes:

Does "cime" mean the highest part of | a mountain ?
| a house ?
| a man ?

or selecting as decoys the words having +A, +B, +C and +D, or E or F:

Select the right pairs | cime . . house
| tête . . mountain
| toit . . man

2. Syntactical structures

Understanding a sentence requires complex operations. No theoretical model exists so far, which helps explaining the performance of a student during the recognition process. How do we describe the process allowing to track the deep structure of the sentence from the parsing of the surface structure, and to associate a semantic meaning to the deep structure ? It certainly includes:

a) A "syntactical competence" (knowledge of the syntagmatical and transformational rules).

b) A "semantical competence" (knowledge of the semantical rules allowing to associate a meaning to the deep structures).

c) The knowledge of the lexicon (phonological and graphical information, syntactical and semantical selection rules associated to each item).

d) Analysis strategies which test (from a, b and c) the most probable hypothesis and discard them as soon as something improper happens. This would imply the memorization of the constituents which have already been analysed. Such processes have been described in a sophisticated way by C.S. Osgood (1).

e) Compensatory devices which supply the deficiencies of the missing information and allow to assess a sentence its most probable meaning even if the sentence is not well-formed or if the message is not completely received.

---

(1) C.S. OSGOOD, 1963.

The ideal procedure would be to analyse and test each of these skills separately. Such a procedure is however unmanageable, so far, because any reading performance involves them all at the same time. However, each of these variables can be approached in some way:

- Vocabulary tests may be conceived for evaluating as accurately as possible the lexical variable. The semantic properties of the words have been described in the previous section. Items have to be written for testing the syntactical properties of the words.
- Morphological tests might be conceived for evaluating the knowledge of the morpho-phonological transformations (for instance, by the use of meaningless words or sentences):

  Ex.: Which is the incorrect sentence ?

  a) des glops paturent

  x b) des glops pature

  Ex.: Which are the plural phrases ?

  a) un patur

  b) un gromp

  x c) des glops

  x d) des paturs

  e) un glop

  x f) des gromps

- Once the researcher has assessed these two dimensions of the student's performance (lexical and morphological skills), he can measure other achievements. By asking questions about kernel sentences (NP-VP), one tests the ability to associate a meaning to a deep structure.

---

x : target response.

It seems, nevertheless, much more difficult to isolate the other variables (knowledge of transformational rules, except morphological ones). When a child assigns a correct meaning to the sentence: "Marie loves Paul" but an incorrect one to "Paul is loved by Marie" (this sentence being understood as "Paul loves Marie"), we can explain this fact:

a) Either because of an imperfect knowledge of the passive transformational rules.

b) Or because the passive transformational rules have not been applied, i.e. the first nominal phrase has been considered as a subject NP and this incorrect hypothesis has not been subsequently discarded.

It should be possible to create tests which would measure separately both of these aspects, but each transformation or set of transformations would require a specific instrument. We are not likely to set up a convenient procedure for generating items devised for testing any transformational rule.

Therefore, it will probably be impossible to test the understanding of transformed sentences by using the deep structure in the question. Items derived this way (1) would measure mixed operative processes. A more exhaustive study applied to well-defined linguistic data would allow to define:

- items measuring the knowledge of the rules.
- items testing the recognition strategies.

---

(1) This procedure has been suggested by J. Bormuth (1969 and 1970).

3. Anaphorical relations

In the two latter sections, we have suggested convenient procedures for testing the knowledge of the vocabulary and of the syntactical structures. It seems much more difficult or even impossible, so far, to test the domain of semantics. Semantical selection rules are not well-known and thus very difficult to test. No criterion tests can here be designed; the construction of tests related to the area of semantics is left this way to the decision of the item-writer.

As mentionned by J. Bormuth , anaphorical relations can however be tested. Bormuth defines an anaphora as a "pronoun-like structure in that it includes both a pro element and an antecedent". In fact, pronouns are one type of anaphora. Anaphora generally serve the function of allowing authors to state a complexly modified concept, set it equal to some shorter form, and thereafter refer to the complex concept using just that shortened form. In order for a reader to understand a discourse, it seems necessary for him to have acquired some set of processes which enable him to identify anaphoric expressions of various types and correctly associate these anaphoric expressions with their proper antecedents.

In this prospect, Menzel's classification of anaphores (1) can be used:

1. Pro Anaphora
2. Referential Repetition Anaphora
3. Formal Repetition Anaphora
4. Class Inclusive Anaphora
5. Synonymous Anaphora
6. Arithmetic Anaphora
7. Inclusive Anaphora
8. Derivational Anaphora
9. Major Anaphora
10. Minor Anaphora

Anaphores do not cover the whole domain of semantics; however, they can supply some evidence about what happens in this area.

Some procedures of derivation of items from anaphorical relations have been described very precisely by J. Bormuth (op. cit. pp. 50-53). It must be emphasized that we want to enlarge the population of items derived from a passage by using the semic schema of words described above (p. 17). This way, we approach at the same time both lexical semantics and inter-santences relations. We are convinced that such a procedure takes into account some of the highest hierarchical skills of the reading process mastered by the students in the lower grades.

---

(1) P. MENZEL, 1970.

By using this method, we have derived items upon anephorical relations from first grade passages. According to the objectives of the test-writer who wants to test this specific skill, it could, however, reveal necessary to write passages containing different kinds of anaphores and to control other variables.

4. Cloze-tests

We have observed in the previous sections that some compensatory devices must be available which supply the deficiencies of the missing information.

Let us have a look at some procedures convenient for testing this skill. Deleting or stumping letters and syllables within words could be used for testing the decoding skills and the word attack processes. Criterion procedures can easily be designed for achieving this goal by selecting derivation rules of items.

Let us now describe four procedures intended to measure reading comprehension which have been successfully used with first and second grade students.

a) From sentences isolated from the context, a word is randomly selected. An item is then written by choosing decoys from a basic vocabulary list. By taking or not syntactical constraints into account, different skills may be approached.

Ex. The item

"Le pipe de * Paul / vite / boit fume"

measures the knowledge of a syntactical constraint.

Ex. The item

"Monique lave son * orangeade / genou / rhume sale"

measures the knowledge of a semantical constraint.

Such items have been devised and administered by E. Boxus and M. Detheux (1) at the Laboratoire de Pédagogie expérimentale of the University of Liège.

b) From sentences isolated from the context, a word is randomly selected and deleted. Such cloze items have been administered by P. DYCKES (2) at the Institut Pédagogique National of Luxemburg. In the same study, he administered other reading sub-tests; by processing the correlation matrix he proved that this skill was at the top of the hierarchy when other sub-tests were considered.

c) From a passage, it proves very easy to build items as those described in a).

d) Cloze tests may be constructed from a first grade passage. A pre-test administered by J.P. Rapaille (3) at the beginning of the schoolyear (2nd grade) proved to be very interesting for it allowed to ascertain that some students are able to use their lexical, syntactical and semantical competence for completing the test. The very skewed distribution of the results reveals, however, that such complex processes are not yet mastered by a majority of students belonging to the age group.

---

(1) Research project in progress.

(2) I.P.N., Luxemburg, 1970.

(3) J.P. RAPAILLE, Research project in progress, Laboratoire de Pédagogie expérimentale, University of Liège.

CONCLUSIONS

As a conclusion, let us now explain which goals could be best achieved through the use of the model suggested.

First of all, the class-room teacher who has to measure day by day the reading achievement while learning is going on. He cannot be satisfied with usual standardized tests ranking students according to a gaussian reference distribution: a teacher who is aware of the real goals of his teaching, cannot be highly concerned by the traditional summative evaluation. What he wants is to know accurately which reading skills have been mastered by a student. Such an accurate evaluation is needed before resorting to remedial procedures. The teacher urgently needs tools to test any reading skill he wants to. It would therefore be of a considerable help if we could provide him with a complete list of possible items. The model suggested can serve that purpose.

In this perspective, we do not have to bother about the difficulty level or the discrimination power of items. All we need to know is that they have been generated by using the model and are representative of the whole population of items measuring the skill. Of course, the teacher is, in most cases, not able to master such a sophisticated technology. Specialized research departments must help.

Educational reading researchers are concerned, too, by this model. It might, for instance, prove very important to refer to valid criterion measures for assessing the predictive power of reading readiness tests or for studying the prerequisites of reading learning. The same

criterion measures might also prove important for identifying learning effects of different reading methods ("look-say", "phonics",...), for calibrating basic readers, etc.

At any level of responsibility, educational authorities are anxious to provide the best education according to the available resources. The better primary reading achievement, the better would be any further school achievement. A reading evaluation model is indispensable for achieving any local or national surveys on reading, for it provides a large population of representative items belonging to well-defined sub-classes and allows by the same token to utilize very reliable item-sampling techniques.

BIBLIOGRAPHY

R. BAILLY, Dictionnaire des synonymes, Larousse, Paris, 1970.

J. BORMUTH, On the Theory of Achievement Test Items, University of Chicago Press, 1970.

- , Development of Readability Analysis, Office of Education, Final Report, Project No. 7.0052, March 1969.

J. CHALL, Learning to Read, The great Debate, New York, Mc Graw Hill, 1967.

E.B. COLEMAN, Collecting a Data Base for a Reading Technology, in Journal of Educational Psychology, Vol. 67, August 1970, pp. 1-23.

A.D. DICK, Stimulus Variables in Single-Letter Recognition, in Psychonomic Science, 20, No. 4, 195, August 1970, II C.3.

J. DUBOIS, R. LAGANE, G. NIOBEY et al., Dictionnaire du français contemporain, Larousse, Paris, 1967.

G. GOUGENHEIM, P. RIVENC, R. MICHEA, A. SAUVAGEOT, L'élaboration du français fondamental (1er degré), Didier, Paris, 1967.

Institut Pédagogique National du Luxembourg, BIM Reading Test, 1970.

P. MENZEL, The Linguistic Bases of The Language of Writing Items for Instruction Stated in Natural Language, in J. BORMUTH, On the Theory of Achievement Test Items, University of Chicago Press, 1970.

O.S. OSGOOD, On Understanding and Creating Sentences, in American Psychologist, 1963, 18, pp. 735-751.

H.M. POPP, Visual Discrimination of Alphabet Letters, in The Reading Teacher, 1964, 17, pp. 221-225.

R. THIMONNIER, Le système graphique du français, Plon, Paris, 1967.